Berger in retrospect: effect of varied weight training programmes on strength

R N Carpinelli

The evidence that multiple sets of exercise are superior to a single set for maximal strength gains, as suggested by Berger in 1962, is reviewed. The validity and practical significance of Berger’s strength training study are questioned. Well controlled, methodologically sound studies that minimise confounding variables are required to support the hypothesis that multiple sets of exercise elicit superior gains in strength.

There is a pervasive misconception that multiple sets of a strength training exercise are superior to a single set for increasing muscular size and strength. The prevalent recommendation, which appears in exercise physiology textbooks and strength training reviews, is to perform multiple sets (at least three) of each exercise. However, there is very little evidence to support the preponderant belief that multiple sets are superior to a single set. Fig 1 illustrates the labyrinth of cross referencing by physiologists who recommend the execution of multiple sets. The only evidence cited to support their multiple set training philosophy is one strength training study by Berger, which is perhaps the genesis of the unsubstantiated belief that multiple sets are required for optimal gains in strength. Therefore, a close examination of this study is warranted. This retrospective is based on the original information published in Berger’s study, as well as personal written and telephone correspondence (1998) with Dr Berger in which he clarified specific methodologies that were not elucidated in the original publication. He expressed his opinion—both before his research and currently—that multiple sets are required for maximal increases in muscular strength and hypertrophy (personal communication; 1998).

BERGER (1962)
Berger reported that nine groups of college age men (about 20 in each group) performed strength training exercises three times a week for 12 weeks. The students (n = 177) were enrolled in one of nine weight training classes and not equated or randomised before training. There was no control group. In addition to their regular weight training programme, which is not described in the report, participants performed different combinations of sets and repetitions in the free weight bench press exercise—for example, three sets of six repetitions (III-6). Berger used roman and arabic numerals to represent sets and repetitions respectively. All nine groups showed a significant increase (p<0.05) in one repetition maximum (1RM) bench press (table 1).

Berger’s Table 4 (p 176) had three previously unchallenged typographical errors. The errors have been corrected (row 6, column 5 changed from III-2 to 1-6; row 2, column 9 changed from III-6 to Groups; and row 3, column 9 changed from Groups to III-6), and the results for the end of 12 weeks of training are presented here in table 2 (personal communication; 1998). Berger notes that groups differing at the 0.05 level are separated by dashes. Table 2 shows which of Berger’s nine groups (36 comparisons) differ significantly and those that do not. After 12 weeks of training, group III-6 is separated by dashes from II-6, but not from 1-6. Similarly, group III-2 is separated by dashes from II-2 but not from 1-2. That is, the increase in 1RM bench press for the III-6 group was significantly greater (p<0.05) than for the II-6 group, but not significantly greater than for the I-6 group. The II-6 group is not separated by dashes from the I-6 group. Likewise, the increase in strength for group III-2 was significantly greater (p<0.05) than for the II-2 group but not significantly greater than for the I-2 group. The II-2 group is not separated by dashes from the I-2 group. If a greater number of sets were really better than fewer sets, three sets would be better than one set, as well as better than two sets—they were not. There was no significant difference in 1RM between groups I-10 and II-10, groups I-10 and III-10, or groups II-10 and III-10. Table 3 shows that seven out of nine of Berger’s comparisons between groups who performed the same number of repetitions (2, 6, or 10) showed no significant difference in the magnitude of strength gains as a result of performing single or multiple sets. Whenever multiple comparison testing is used, the chance of making a type I error is increased—that is, rejecting the null hypothesis when the null hypothesis is actually true. Perhaps the differences that Berger reported were due to chance or random error (personal communication; 1998).

Berger also compared the results of training with one, two, or three sets by combining the nine groups according to the number of sets performed. Training with either one set or two sets produced similar improvements in strength (22.3% and 22.0% respectively), whereas training with three sets elicited an increase of 25.3%. The difference (3.2%) between training with one set and three sets was significant (p<0.05) after 12 weeks of training. If it is assumed that the difference was not caused by random error, the practical significance of a 1.8 kg difference as a result of performing three sets compared with one set...
Berger's study (1958) is questionable (personal communication; 1998). That is, for most of the strength training population and with the exception of elite strength athletes, a 3% difference in strength may not be worth 300% greater time and energy expenditure.

Berger reported that the greatest difference in the rate of improvement between the combined one set, two set, and three set groups occurred during the second three weeks (weeks 4–6) of training. During this period the increase was 5.0%, 5.7%, and 7.8%, for the one, two, and three set groups respectively. Perhaps there was a greater learning effect for the combined three set groups in the early phase of the study. A movement such as the free weight bench press exercise, which requires a high level of strength, is considered a skilled act, and control of the involved muscles by the nervous system is very complex. Therefore, when an unfamiliar exercise is introduced, the increase in strength is largely because of adaptations in the nervous system that optimise control of the muscles involved in the exercise and the coordination necessary to perform the exercise efficiently. These neural adaptations are predominant in the first few weeks of training.

For the rest of the 12 week study (weeks 7–12), the difference between groups gradually diminished until the rate of improvement was practically the same during the last three weeks of training. In fact, the rate of improvement for weeks 9–12 was 4.0%, 3.2%, and 3.3% for the one, two, and three set groups respectively. Similarly, for the specific number of repetitions (six) that Berger claimed to be optimal, the rate of strength increase for the last three weeks was 4.8%, 4.0%, and 4.0% for group I-6, II-6, and III-6 respectively.

Table 1

<table>
<thead>
<tr>
<th>Group (n)</th>
<th>↑ kg</th>
<th>% ↑</th>
</tr>
</thead>
<tbody>
<tr>
<td>II-2 (18)</td>
<td>11.3</td>
<td>20.0</td>
</tr>
<tr>
<td>III-2 (18)</td>
<td>9.3</td>
<td>17.3</td>
</tr>
<tr>
<td>IV-2 (18)</td>
<td>13.3</td>
<td>23.5</td>
</tr>
<tr>
<td>VI-2 (20)</td>
<td>14.5</td>
<td>25.5</td>
</tr>
<tr>
<td>VI-6 (21)</td>
<td>16.7</td>
<td>29.9</td>
</tr>
<tr>
<td>VI-10 (19)</td>
<td>12.2</td>
<td>21.6</td>
</tr>
<tr>
<td>III-10 (19)</td>
<td>14.2</td>
<td>25.1</td>
</tr>
<tr>
<td>IV-10 (19)</td>
<td>13.0</td>
<td>23.0</td>
</tr>
</tbody>
</table>

Data from Berger.

Table 2

<table>
<thead>
<tr>
<th>Groups</th>
<th>Mean (kg)</th>
<th>Significant difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>III-6</td>
<td>73.3</td>
<td>III-6</td>
</tr>
<tr>
<td>I-6</td>
<td>71.0</td>
<td>I-6</td>
</tr>
<tr>
<td>I-10</td>
<td>70.8</td>
<td>I-10</td>
</tr>
<tr>
<td>II-30</td>
<td>69.6</td>
<td>II-30</td>
</tr>
<tr>
<td>III-10</td>
<td>69.6</td>
<td>III-10</td>
</tr>
<tr>
<td>III-6</td>
<td>69.5</td>
<td>III-6</td>
</tr>
<tr>
<td>I-10</td>
<td>68.8</td>
<td>I-10</td>
</tr>
<tr>
<td>I-2</td>
<td>67.9</td>
<td>I-2</td>
</tr>
<tr>
<td>II-2</td>
<td>66.4</td>
<td>II-2</td>
</tr>
</tbody>
</table>

Data from Berger. "Groups differing at the 0.05 level are separated by dashes" (Berger, 1962, p.176).
They all performed and were graded on other exercises besides in combination with two, six, or ten repetitions. The whether one, two, or three sets might be better systematically significant interaction between sets and repetitions—that is, 9, or 12 weeks). In Conclusion 6, Berger interaction was not significant at any period of training (3, 6, or 12 weeks). In Conclusion 6, Berger 

greatest volume of exercise—is the most effective. However, contrary to Berger's claim that the combination of three sets and six repetitions, compared with subjects in the III-10 group (23.0%), who performed 30 repetitions per session, the III-10 group (23.0%), who performed 30 repetitions per session. Berger used an analysis of covariance to test for any significant interaction between sets and repetitions—that is, whether one, two, or three sets might be better systematically in combination with two, six, or ten repetitions. The F ratio for interaction was not significant at any period of training (3, 6, 9, or 12 weeks). In Conclusion 6, Berger contends that training with one, two, or three sets in discrete combination with two, six, or ten repetitions (interaction) was not systematically more effective in improving strength than other combinations. However, in Conclusion 7, he declares that a combination of six repetitions and three sets was more effective in improving strength than any other combination of sets and repetitions. Berger reports no significant interaction, and then antithetically claims that the combination of three sets and six repetitions is more effective than any other combination of sets and repetitions.

Berger’s subjects were enrolled in weight training classes. They all performed and were graded on other exercises besides the bench press (personal communication; 1998). However, Berger reported only the results of the bench press exercise. He did not control the number of sets and repetitions performed for the other weight training exercises (personal communication; 1998), and those exercises involving the triceps, pectorals, and deltoids may have confounded the results of the 1RM bench press. Participants chose their own rest time between sets and exercises (personal communication; 1998). Although the optimal rest time between sets or exercises is unknown, similar rest between sets would have eliminated recovery time as a confounding variable. Berger did not control for the speed of movement, which was self selected by each participant (personal communication; 1998). Different speeds of movement within and among the nine groups may have affected his results.

Participants were given enough assistance to accomplish a specific number of repetitions (2, 6, or 10). The amount of assistance, the specific spotter for each participant, the number of assisted repetitions, and how often the assisted repetitions were applied in each group for each participant were neither controlled nor documented (personal communication; 1998). When a subject thought that he could perform more than the designated number of repetitions, he terminated the set at the designated number and increased the weight at the next session (personal communication; 1998)—that is, participants did not exceed the designated number of repetitions (2, 6, or 10). Completing every set to muscular fatigue would have ensured similar maximal exercise intensity (effort) for everyone. Because the sets were terminated at a specific number of repetitions, there was no control for exercise intensity (degree of effort or fatigue) within or among the groups. The concept of muscular fatigue is highly subjective and not universally defined or fully understood. As fatigue begins to develop within the muscle as a result of several sub-maximal muscle actions, more motor units are recruited and motor unit firing rates increase in an attempt to maintain the desired force output. When the force generating capacity drops below the required force, despite full recruitment and optimal firing rates of motor units, it is generally defined as muscular fatigue. The causes may range from motivation to maintain voluntary effort to failure in the contractile mechanism. The various processes that cause fatigue during high intensity strength training may provide the stimulus for increased strength. It is speculated that the degree of motor unit activation determines the magnitude of the strength training response and that intensity is the most contributing factor for increasing strength.

### Table 3

Comparison of groups performing a similar number of repetitions [2, 6, or 10] after 12 weeks

<table>
<thead>
<tr>
<th>Groups</th>
<th>3RM</th>
</tr>
</thead>
<tbody>
<tr>
<td>III-6 and II-6</td>
<td>SD</td>
</tr>
<tr>
<td>III-6 and I-6</td>
<td>NSD</td>
</tr>
<tr>
<td>II-6 and I-6</td>
<td>NSD</td>
</tr>
<tr>
<td>III-2 and II-2</td>
<td>SD</td>
</tr>
<tr>
<td>III-2 and I-2</td>
<td>NSD</td>
</tr>
<tr>
<td>II-2 and I-2</td>
<td>NSD</td>
</tr>
<tr>
<td>III-10 and I-10</td>
<td>NSD</td>
</tr>
<tr>
<td>III-10 and I-10</td>
<td>NSD</td>
</tr>
<tr>
<td>II-10 and I-10</td>
<td>NSD</td>
</tr>
</tbody>
</table>

Data from Berger. SD, Statistically significant difference (p<0.05) between groups in one repetition maximum (1RM) bench press; NSD, no statistically significant difference between groups.

In the Results section, Berger claims that more sets, more repetitions per set, and more total repetitions at each training session resulted in a greater improvement in strength. His implication is that three sets of ten repetitions (III-10)—the greatest volume of exercise—is the most effective. However, contrary to Berger’s claim, table 2 shows that there was a significantly (p<0.05) greater increase in strength for subjects in the III-6 group (29.6%), who performed 18 repetitions per session, compared with subjects in the III-10 group (23.0%), who performed 30 repetitions per session.

Berger used an analysis of covariance to test for any significant interaction between sets and repetitions—that is, whether one, two, or three sets might be better systematically in combination with two, six, or ten repetitions. The F ratio for interaction was not significant at any period of training (3, 6, 9, or 12 weeks). In Conclusion 6, Berger contends that training with one, two, or three sets in discrete combination with two, six, or ten repetitions (interaction) was not systematically more effective in improving strength than other combinations. However, in Conclusion 7, he declares that a combination of six repetitions and three sets was more effective in improving strength than any other combination of sets and repetitions. Berger reports no significant interaction, and then antithetically claims that the combination of three sets and six repetitions is more effective than any other combination of sets and repetitions.

Berger’s subjects were enrolled in weight training classes. They all performed and were graded on other exercises besides the bench press (personal communication; 1998). However, Berger reported only the results of the bench press exercise. He did not control the number of sets and repetitions performed for the other weight training exercises (personal communication; 1998), and those exercises involving the triceps, pectorals, and deltoids may have confounded the results of the 1RM bench press. Participants chose their own rest time between sets and exercises (personal communication; 1998). Although the optimal rest time between sets or exercises is unknown, similar rest between sets would have eliminated recovery time as a confounding variable. Berger did not control for the speed of movement, which was self selected by each participant (personal communication; 1998). Different speeds of movement within and among the nine groups may have affected his results.

Participants were given enough assistance to accomplish a specific number of repetitions (2, 6, or 10). The amount of assistance, the specific spotter for each participant, the number of assisted repetitions, and how often the assisted repetitions were applied in each group for each participant were neither controlled nor documented (personal communication; 1998). When a subject thought that he could perform more than the designated number of repetitions, he terminated the set at the designated number and increased the weight at the next session (personal communication; 1998)—that is, participants did not exceed the designated number of repetitions (2, 6, or 10). Completing every set to muscular fatigue would have ensured similar maximal exercise intensity (effort) for everyone. Because the sets were terminated at a specific number of repetitions, there was no control for exercise intensity (degree of effort or fatigue) within or among the groups. The concept of muscular fatigue is highly subjective and not universally defined or fully understood. As fatigue begins to develop within the muscle as a result of several sub-maximal muscle actions, more motor units are recruited and motor unit firing rates increase in an attempt to maintain the desired force output. When the force generating capacity drops below the required force, despite full recruitment and optimal firing rates of motor units, it is generally defined as muscular fatigue. The causes may range from motivation to maintain voluntary effort to failure in the contractile mechanism. The various processes that cause fatigue during high intensity strength training may provide the stimulus for increased strength. It is speculated that the degree of motor unit activation determines the magnitude of the strength training response and that intensity is the most contributing factor for increasing strength.

### BERGER (1963)

If three sets of six repetitions (3 × 6) were superior to other training protocols, subsequent studies would have replicated those results. However, a follow up study by Berger failed to support his conclusion that 3 × 6 is the best training protocol. He trained three groups of young men three times a week for nine weeks performing a 3 × 6 (n = 14), 3 × 10 (n = 19), or 6 × 2 (n = 15) protocol in the free weight bench press exercise. All groups showed a significant (p<0.001) increase in 1RM bench press (21.3%, 20.0%, and 16.9% respectively), with no significant difference among the groups. That is, the 3 × 6 protocol was not superior to the 3 × 10 protocol. In Berger’s previous investigation, he reported that the 3 × 6 protocol showed a significantly greater increase in 1RM bench press than the 3 × 10 protocol. Berger concluded that further research was required to determine the optimum combination of sets and repetitions for the greatest improvement of strength. In a summary of his strength training research, Berger recommends executing at least three sets at each session. He cites his 1962 study in support of his high volume training philosophy.

### BERGER (1972)

Berger trained three groups of young men three times a week for eight weeks. They performed one (n = 57), two (n = 20), or three sets (n = 20) of static bench press exercise by attempting to raise an immovable bar with a maximum effort for 6–8 seconds at two positions (with the bar on the chest and 90° elbow flexion). There was a significant (p<0.01) increase in dynamic 1RM bench press strength for the one, two, and three set groups (19.8%, 15.4%, and 21.7% respectively), with no significant difference among the groups. Berger concluded that training with one, two, or three sets produces approximately the same improvement in dynamic strength. However, he subsequently recommended performing at least three to five sets, substantiated only by his 1962 study.

### DISCUSSION

“There is very little peer reviewed evidence to support a high volume strength training protocol.”

In his book Conditioning for men, Berger recommends performing three to five sets of each exercise for optimal strength, and in his Introduction to weight training, he...
recommend performing five to nine sets of each exercise. In his textbook *Applied exercise physiology*, Berger claims that three sets are more effective than fewer sets for maximising strength. The only reference that Berger cites in the aforementioned books is his 1962 strength training study. Berger still believes that a great volume of exercise is required for optimal strength gains (personal communication; 1998). However, there is very little peer reviewed evidence to support this theory, as decades of research have consistently shown that a minimum of three sets of each exercise is required to produce adaptations or acute responses to resistance exercise. For example, in a well controlled study, Starkey et al. 14 randomly assigned healthy untrained men and women (18–50 years of age) to one set (n = 18), three set (n = 20), or control (n = 10) group. Both training groups performed 8–12 repetitions to volitional fatigue using a two second concentric and four second eccentric protocol for dynamic bilateral knee extension and knee flexion exercises on two MedX strength machines three times a week for 13 weeks. They did not participate in any other form of training for the duration of the study. The three set group rested for one to three minutes between sets. Resistance was increased about 5% at the subsequent session when 12 or more repetitions were completed in good form. All sessions were monitored by the investigators, and the amount of resistance, number of repetitions, and rating of perceived exertion were recorded after each set. Peak knee extension torque at seven angles significantly increased (p<0.05) in both the one set (18–36%) and three set (13–27%) groups. Peak knee flexion torque at the same angles increased (p<0.05) in both the one set (13–35%) and three set (8–41%) groups. There was no significant difference between the one set and three set groups in either knee extension or knee flexion peak torque. Both groups had similar significant increases (p<0.01) in dynamic training resistance for both exercises. Ultrasound scans showed a significant increase (p<0.01) in muscle thickness, with no significant difference between the one set and three set groups.

Hass et al. 15 randomly assigned 42 male and female recreational weightlifters (20–50 years of age), with an average of 6.2 years of strength training experience, to one of two training groups: one set (n = 21) or three set (n = 21). All participants were previously using the one set protocol (one circuit of nine exercises) three times a week. The one set group continued to perform one set of each exercise and the three set group performed three sets of each exercise on the nine MedX machines three times a week for 13 weeks. They executed each set of 8–12 repetitions using a two second concentric and four second eccentric protocol to volitional fatigue. Subjects rested one minute between exercises and three to five minutes between circuits. The investigators monitored each training session and recorded the resistance, number of repetitions, and rating of perceived exertion after each exercise. When 12 or more repetitions were completed in strict form, the resistance was increased 5–10% at the next session. Dynamic muscular strength was assessed using the 1RM on five of the nine MedX machines, as well as maximal isometric knee extension and knee flexion strength on a MedX ergometer. Seven skinfolds and six circumferences were assessed to estimate body composition. After 13 weeks, both groups showed a significant increase (p<0.05) in dynamic 1RM strength (8–14%) on the five exercises, with no significant difference in the magnitude of strength between the one set and three set groups. Both groups significantly increased (p<0.05) maximal isometric knee extension and knee flexion torque, with no significant difference between groups. Lean body mass significantly increased (p<0.05) in both groups, with no significant difference between groups. There were originally 49 subjects in this strength training study, but five had to be removed by the investigators because of poor compliance, and two other participants withdrew because of injuries. All seven subjects who did not complete the study were from the three set group. The greater volume of exercise (multiple sets) in the three set group may have contributed to their lower compliance.

### CONCLUSION

One consequence of accepting an unsubstantiated training philosophy is that many researchers have assumed for decades that a minimum of three sets of each exercise is required to produce adaptations or acute responses to resistance exercise. For example, as a result of different strength training programmes, adaptations have recently been reported for women, men, and both sexes, as well as acute responses in women, men, and both sexes. None of the researchers attempted to justify their exercise protocols, which comprised at least three sets of each exercise. Perhaps the greater ramification is that many dedicated trainees spend excessive time and perform an unnecessary volume of exercise (three to five sets compared with one set).
The genesis of the belief that multiple sets of each exercise are superior to a single set for maximal strength gains is one very poorly controlled 40 year old strength training study by Berger. The evidence to support the performance of multiple sets is extremely weak. Most of the evidence suggests that single and multiple sets produce similar increases in strength.

"More research is required to determine if the potential health benefits from strength training require multiple set protocols."

More research is required to determine if the potential health benefits from strength training require multiple set protocols. However, Kelley and Kelley in a meta-analysis of randomised controlled trials reported that various strength training protocols, such as the number of sets, did not influence the significant decrease in resting systolic and diastolic blood pressure in adults.

Replication of Berger’s results, using greater controls and randomisation, and a corroboration of evidence from different researchers is necessary to support the practical significance of performing multiple sets of each exercise. If the experimental treatment—for example, the number of sets—had a really strong effect on the outcome—for example, muscular strength or hypertrophy—it would influence the outcome whenever the experiment, or a similar investigation, is repeated. It should be recognised that no well controlled study has shown single set training to be superior to multiple set training—that is, there appears to be no detrimental effect from performing multiple sets. The benefit of single set training is time efficiency because most of the time it elicits similar strength gains in less time. Only five studies report a significantly greater increase in strength as a result of multiple sets. The preponderance of research reports no difference in the magnitude of strength gains or muscular hypertrophy as a result of performing a greater number of sets.

REFERENCES
35 Westcott WL. 4 key factors in building a strength program. Scholastic Coach 1986;65:104–5, 123.


